INTERACTION PATTERNS FOR AUDITORY USER INTERFACES

C. Frauenberger, V. Putz, R. Höldrich, T. Stockman

University of Music and Dramatic Arts Graz, Institute of Electronic Music and Acoustics
Queen Mary University of London, Department of Computer Science
Graz University of Technology, Signal Processing and Speech Communication Laboratory
frauenberge@dcs.qmul.ac.uk, putz@iem.at, hoeldrich@iem.at, tonys@dcs.qmul.ac.uk

ABSTRACT

This paper proposes the use of interaction patterns in the design process of auditory displays in human-computer interaction. To avoid introducing visual concepts in auditory design, a common ground for developing user interfaces without determining their means of representation is proposed. This meta domain allows for the design of user interfaces which can be equally realised in different interaction modalities or multi-modal settings. Although this work focuses on the auditory domain the concept shown is developed keeping in mind that it should be equally applicable in other modalities. A set of mode independent interaction patterns for design in the meta domain are introduced along with their transformation into the auditory domain. A real world application was chosen to evaluate the approach. MS Explorer was analysed, described through the mode independent interaction patterns and transformed into the auditory domain making extensive use of 3D audio rendering techniques. The result, a virtual audio reality version of a file manager, was evaluated with normally sighted persons as well as visually impaired and blind participants showing the feasibility and usability of the approach.

1. INTRODUCTION

When designing auditory displays for user interfaces, designers can not count on commonly accepted design principles or guidelines for robust end efficient GUI interfaces. In contrast to the visual domain the auditory domain is heavily underused and development has not yet addressed guidelines to an extent that enables user interface designers to create auditory displays without much expertise or experience. This paper proposes an approach towards establishing methods to make the design of effective auditory user interfaces easier and more straight forward. We propose the use of interaction patterns in the design process and will subsequently elaborate their development and utilisation in a prototype design of an auditory file manager application.

A major difficulty when designing user interfaces for alternative interaction channels is assuming that visual concepts transfer across to other domains. Thinking visually when designing auditory displays leads to bad designs, because the properties of human senses are substantially different. This challenge is addressed in our approach by introducing a mode independent meta domain, in which user interfaces may be designed without the determination of their means of representation. User interfaces can then be realised from their mode independent description, maintaining the strengths of the interaction channel. Although this work is focused on the realisation in the auditory domain, we keep in mind that the concept should be equally applicable in other modalities.

The following sections highlight the motivation behind the approach, the state-of-the-art of auditory design, and what our objectives are. Subsequently, section 2 describes the approach of using interaction patterns for the design of user interfaces, describing the idea, explaining the patterns developed and discussing a detailed example. Section 3 describes the usability evaluations conducted, including the details of the test design, the evaluation and analysis. Finally, we draw conclusions and state ideas for future work.

1.1. Motivation

Audio has been part of user interfaces since the very early days of computers. Sounds were first used to indicate errors or warnings when processing requests of the user. Sound was an appropriate way to alert users and the simple beeps used in such interfaces were about the same level of sophistication as the graphical output at that time. Since then, the development of graphical user interfaces has been substantial, greatly outstripping the development effort and understanding of auditory interfaces. Currently audio is radically under used in human-computer interfaces, being still predominantly used to signal the occurrence of an error, or to draw the user's attention to the occurrence of a specific event. Sound remains largely unused as a fully integrated element of multimedia interfaces, and is little used in support of mainstream task performance.

Important reasons to seek to change this in the context of modern user interface design can be summarised by [1, 2]:

- Increased complexity of tasks
- Miniaturisation of devices
- Mobility of the user
- Naturalness of interfaces
- Accessibility for disabled users

Many requirements for an improved user interface resulting from the points above can be achieved by auditory displays, whether sound is being used as an element of a multimedia interface, or on its own as a single medium. To prevent user interfaces from being overloaded due to the complexity of tasks, it is necessary to utilise multi-modal approaches to exploit human sensory capabilities to their optimum. Auditory displays can be designed as virtual or augmented reality and do not require much physical space. Sound is also highly portable and integrable in mobile devices. Since audio is an important part of our every-day environment, the integration of audio clearly improves the naturalness in human-computer interaction. Finally, visually impaired users rely on non-visual modes to interact with the increasing range of information technologies integrated in our lives.

Past research has shown that audio is capable of contributing much more than beeps to user interfaces. There are, however, significant problems in understanding human methods of acoustic communication and how complex tasks in human-computer interaction can be rendered in the auditory domain. Many prototypes have been developed and evaluated and satisfactory solutions have been created for specific applications. However, as audio displays become more widespread, user interface designers must be provided with robust, but customisable patterns and design principles.

1.2. State of the Art

The increasing computational power available for digital signal processing made increasingly complex simulations of acoustical environments possible. The term *Virtual Audio Environment* describes the simulation of acoustical scenes with sound reproduction techniques. The goal is to create natural environments which are customisable and controllable in real time, very much as visual virtual environments were developed [3]. Acoustical rendering of objects (sound sources), the environment and the listener can be realised using a number of different approaches, such as Ambisonics[4], Wave Field Synthesis [5] or Vector Based Amplitude Panning [6]. Recently these techniques have been utilised to create spatial auditory displays[7, 8].

While these techniques may provide accurate acoustical rendering, there are still many problems with using them in auditory displays. Most of the problems are related to the human capacity for auditory perception. A problem still subject to investigation is sound source segregation in virtual environments when rendering sound sources concurrently [9], another issue is providing robust orientation cues for navigation [10]. Further challenges result from complex psychoacoustic effects such as informational masking, not fully understood in the context of auditory displays [11].

A variety of auditory displays were developed for specific problem domains (e.g.: [12, 13, 14]) and some research done towards a structured approach to more generic solutions. Early proposals include the *Mercator* project, the first framework targeting customary Unix desktops [15, 16]. Another proposal was *Y-Windows* also following the idea of building alternative, audio rendering engines (servers) for existing clients requesting their user interface representation [17]. However, both approaches implied that graphical concepts were translated into the auditory domain and therefore had their limitations. A first attempt to address this problem was made in [2] and subsequently led to the proposal described in this paper.

Former work that investigated well-known HCI concepts in the context of auditory display design include audio metaphors [18] and other structural approaches to include sound into human-computer interaction (Earcons[19]). Recently, the proposal of using patterns in sonification highlights the advantages of such methods in re-usable designs [20, 21]. Related research towards establishing a common model for the design space taking into account all human interaction channels includes [22]. Here the author also stresses that *"While applications that use multi sensory information 'displays' are becoming more common, frameworks to assist in design of these displays need to be developed"*

Auditory display design needs basic heuristics to assess user satisfaction similar to those in the graphical domain [23]. Examples of where work is needed to identify heuristics to guide the process of auditory display design including minimizing the problems incurred due to the transient nature of sound, quantifying how the effectiveness of interactions can be improved through learning and providing guidelines for how sound can best be integrated with other media [24]. Also the design pattern method is a promising approach and other design principles may also give more control over the efficiency of auditory displays.

1.3. Objectives

The translation of graphical user interfaces into the auditory domain reveals many problems related to the general differences of the senses. Many communication principles work entirely differently to one another, and the strengths and weaknesses of the senses are almost in diametric opposition to each other. Key differences between hearing and vision include, for example:

- high spatial resolution vs. low spatial resolution
- cut out and planar projection vs. full 3D surrounding
- high user attention vs. variable user attention
- single focus vs. parallel processing

If not considered in the design, these issues can significantly reduce the effectiveness of auditory displays. Design practice is, however, dominated by visual concepts and interface designers tend to think visually when creating auditory user interfaces.

This paper proposes a different approach to solve this problem. We propose a mode independent description of user interfaces, as common ground for the design of either visual interfaces or auditory interfaces or any mixed multi-modal variants. The principle behind the proposal is to maintain the particular strengths of any interaction modality by transforming the abstract task (e.g. mode independent → audio) instead of an already existing interpretation of the task (e.g. vision → audio)

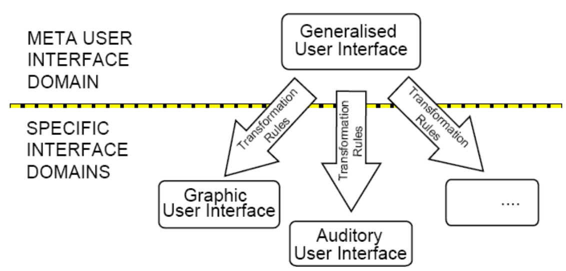

Figure 1: *Mode independent meta domain as common basis*

Figure 1 shows the concept. The goal is to find mode independent interaction patterns to describe the required interface in a meta-domain. Transformations can then be developed to create actual displays. In analogy to the concept of XML stylesheets, the automatisation of this process would make very flexible user interfaces possible. This, however, may not be achieved without knowing much more about the mechanisms of interaction in different modalities.

The following section describes the concept of mode independent interaction patterns, examines their structure and provides an example of how to transform them into the auditory domain.

2. INTERACTION PATTERNS

The concept of patterns is well-known in user interface design, but is very much focused on vision. Furthermore, most existing pattern sets are created from the perspective of the designer. Our approach, which is based on user-centred pattern design led to a set of task related patterns addressing user problems (e.g. the need to be able to locate specific commands and to find out how to activate them), based upon several interaction principles [25]. We chose this set of interaction patterns from [26] as the basis for the development of the mode independent patterns because they were user-centred and already stressed the abstract description of user requirements.

In order to detach the mode from the interaction patterns the following terminology was used: The *representation medium* means the domain or the combination of the domains in which the user interface will be realised. Within this representation medium there are *representation areas* defined which provide the boundary for *objects* of the user interface. These objects may result from one or more *interaction patterns* transformed into the representation medium. Despite these small changes to the terminology, the existing interaction patterns were easy to re-formulate so that they would not prejudice the process of realisation.

Table 1 gives an overview of the patterns developed. The full description would include the *Name*, *User problem*, *Conditions* to consider when using this pattern, a generic *Solution* that indicates how this user problem shall be solved and the *Attributes* to be used to link this pattern in its context. Furthermore, a graphical and auditory example of a realisation is stated with every pattern. However, a full description is out of the scope of this paper, and so table 1 provides a brief description of *Name*, *User problem* and *Acoustic realisation*. The set developed is described in full detail in [27, 28].

2.1. Structure

While developing the patterns, we recognised that certain tasks or parts of patterns recurred in other patterns too. This led to the concept of atoms and contextual attributes. Similar to a vocabulary for instantiating designs, a set of atoms were developed from which patterns may draw when addressing a particular set of user requirements. This also implies consistent representation of similar elementary units throughout the whole interface although atoms are not sufficient to solve any interaction problem. The name indicates that this level should be the smallest piece in any pattern to avoid over exaggerating modularity.

In order Not to end with a totally unrelated patchwork of small pieces of a user interface, each atom provides contextual attributes. These attributes need to be set by the parent pattern in order to indicate their context. In the graphical domain this would, for example, mean that certain elements like buttons or text fields are *in the same window* sharing the same frame and background colour. The following contextual attributes were identified for our set of atoms:

Similarity: Atoms in the same pattern share properties like timbre, rhythm or type of voice in their acoustical representation.

Proximity: Atoms in the same pattern are grouped together based on the available dimensions of the representation area (space or pitch ranges).

Homogeneity: The same types of atoms should be placed adjacently in a pattern on the basis of the available dimensions of the representation area (space or pitch ranges).

It is important to state that not only the patterns and the atoms undergo the transformation process in order to form a real user interface, but also the contextual attributes must be mapped into the different representation media. Their realisation in the auditory domain will differ considerably from the visual domain.

2.2. An Example

In order to show how these patterns are employed, this section describes the *Command Area* pattern in more detail. It is addressed to solve an interaction problem which has already been mentioned previously: The user needs to know where to find the possible commands and how to activate them. When solving this problem, several further conditions have to be taken into account: 1) Immediate access to all available functions of the application increases interaction speed but consumes a large amount of the available representation area. 2) The amount of the remaining representational media space for the main working area should be kept as large as possible. 3) Concurrent representation of many objects increases the cognitive load. 4) Some functions are used more often then other functions. 5) Some functions need additional parameters to be set by the user before they can be executed. The solution describes the creation of a recognisable area in the representation medium which contains shortcuts to important functions of the applications and a menu. The shortcuts exploit affordances to indicate the usage and use metaphors (natural mapping) to save representation space. Atoms used by the pattern are the link atom and the triggering element atom. The atoms will be connected to the pattern by similarity attributes and triggering elements will be grouped together - if necessary - in meaningful sub-groups (e.g. for a menu).

Figure 2 shows a graphical representation of this solution - the command area of a Microsoft Word application. Buttons (including a clear affordance) represent the triggering elements exploiting metaphors (e.g. blank sheet) to indicate the meaning. The area is framed and has the same back-colour. The links at the top provide access to all functions by sub-grouped triggering elements.

Transforming the same interaction pattern to the auditory domain was performed for our evaluation test described in section 3. The two main parts - the menu headers and the toolbar - are situated alongside two different walls of the virtual room as shown in figure 3. The triggering elements of the toolbar are presented

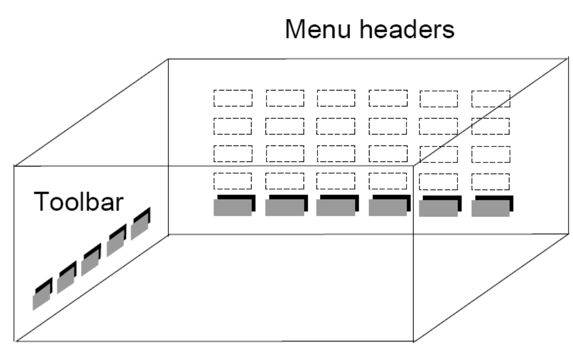

Figure 3: *The layout of the auditory representation of the Command Array pattern*

with short rhythmic patterns (Earcons). The toolbar elements are

Table 1: An overview of the patterns including user tasks and acoustic realisations

Name	**User problem**	**Auditory realisation**
Command Area	The user needs to know where to find the possible commands and how to activate them	A menu area in the 3D space with selectable auditory objects which spatially unfold their contents
Wizard	The user wants to achieve a single goal but several decisions need to be made before the goal can be achieved completely	A sequence of similar acoustical rooms with designated entries and exits
Contextual Menus	At any time users need to know what their possibilities are in order to decide what to do	Sticky objects with trigger items which surround the user and perform the same movements as the user
Mode Cursor	The user is creating or modifying an object and needs to know which edit function is selected	Background sounds indicating a state
Setting Attributes	Users want to see the attributes of the objects they are working on and additionally they need to know how to modify them	Auditory objects have special sound properties attributed indicating the state. Altering the state is shortcut into the command area.
Link in the real world	The user needs to know how to control an object in the interface which resembles an object the user knows from the real world	Special interaction devices, may integrate speech recognition
List browser	The user needs to browse or process several items out of a set or list	Spatially distributed auditory objects with indicated order and browsable with an acoustical lens
Continuous filter	The user needs to find an item in an ordered set	Flexible search fields pre-sorting a list selection
Preview	The user searches one item in a set of items and tries to find it by browsing the set. Often, other search criteria than the items name are more effective	Switching temporarily to a different virtual room with a preview of the content (auditory cartoons)
Navigating between categories	The user needs to access an amount of information which cannot be put on the available representation area	Acoustically themeable groups of content with fast navigation shortcuts
Container navigation	The user needs to find an item in a collection of containers	Bordered space in a 3D environment like Walls or entire Rooms including an intelligent way of navigating through this environment
Unambiguous format	The user needs to supply the application with data but may not know which type or format of data is required or what syntax to use	Represent the properties of an interaction object with common sound properties or with detailed speech instructions on demand
Focus / Selection	Users want to quickly get information about an object they see and possibly want to modify the object	Selected objects change their acoustic properties, e.g. become louder and play in the foreground
Frame / Grouping Layout	The user needs to quickly understand information and its structure. Several different objects have to share a limited amount of representational space	Spatial grouping in the virtual environment e.g. Walls and rooms
Progress	The user wants to know whether or not the operation is still being performed as well as how much longer the user will need to wait	Background sound information exploiting metaphors for progress (like a raising pitch)
Hinting	The user needs to know how to select functions, especially if they are accessible in more than one way (menu, keyboard shortcut, toolbar)	Short spoken hints initiated by a function key
Warning / Message	The user may unintentionally cause a problem situation which needs to be resolved	Sticky objects at the users location which require user interaction
Shield	The user may accidentally select a function that has irreversible effects	Sticky objects at the users location which require user interaction
Preferences	Each user is different and prefers to do things slightly different	A certain room in the virtual environment controlling the options of the application
Favourites	The user needs to find a regularly used item within a large set of items	User defined shortcuts allowing ”to beam” to a certain location within the virtual environment

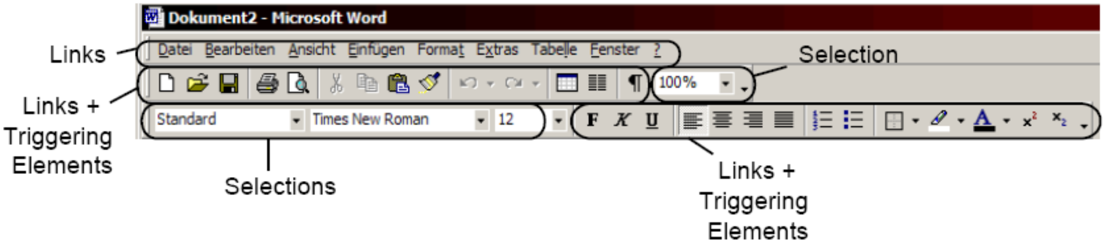

Figure 2: *A Command Area interaction pattern in the graphical domain*

connected by proximity and homogeneity with regard to space and similarity with regard to the sound design. The menu headers are presented as a combination of speech and instrumental sound. The name of each menu item is spoken accompanied by an unobtrusive sound with rising pitch, starting with lowest pitch at the first header item and increasing in pitch while moving towards the last one. The menu headers are connected by proximity and homogeneity with regard to space and pitch and by similarity with regard to the compound sound design (speech and instrumental tone). The sub menu items belonging to the menu headers are placed above the headers within the virtual room, also providing compound sounds.

3. EVALUATION

For evaluation purposes a real world application was chosen and analysed. Its user interface was described through the mode independent interaction patterns and then transformed to the auditory domain. As one of the key applications on a computer desktop a file manager was chosen - the Microsoft Windows Explorer.

3.1. Prototype Design

For testing, the complexity of the application was reduced, rarely used functions were not implemented. The remaining menu structure is shown in figure 4. It still covers the basic functionality of a file manager. From the standard buttons toolbar, five shortcuts

File	Edit	View	Favourites	?
New Folder	Undo	Statusbar	www.iem.at	Information
Delete	Cut	List	www.orf.at	Support
Rename	Copy	Details	www.google.com	
Properties	Paste	Sort		
Close	Select All	Change To		
		Reload		

Figure 4: *Implemented functions of the auditory Explorer version*

were implemented: Undo, Delete, Cut, Copy, Paste. Through the context menu the functions Cut, Copy, Delete, Rename and Properties were available.

The described functions of the Explorer application were analysed and interaction patterns were identified. The *Container Navigation* pattern was used to describe the two main frames of the Explorer. For the folder tree in the left frame the *tree structure* atom was used and the *list* atom described the right content frame. The *Command Area* pattern described the menu structure and the tool bar area. Finally, the *Contextual Menu* pattern solved the availability of the context menu and the *Message* pattern was used for all pop-up windows at their occurrence. Figure 5 shows the basic layout of the virtual audio environment into which the patterns were transformed. The container navigation pattern was transformed

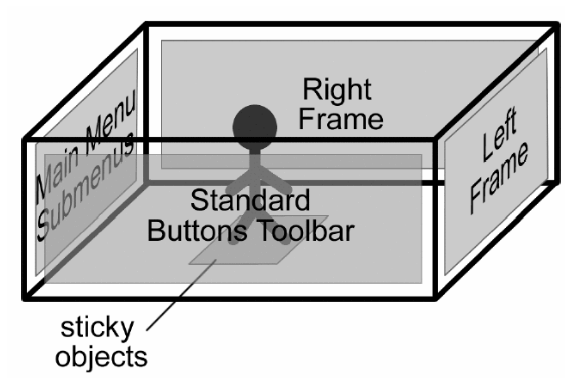

Figure 5: *The virtual audio environment for the Explorer*

into the auditory domain as two different areas in the virtual environment, the walls to the front and to the right. The list at the front wall was stretching out to the top using speech for the name of the items with different voices to indicate the type (folder or file). The tree structure was laid out on a grid on the wall to the right with the left-bottom corner being the root and the right-top corner the last hierarchical level. Unfolding and folding a node in the tree was also implemented. In both areas the user was able to select items and get a context menu. The contextual pattern was realised as sticky objects following the user wherever she moves. The content of the contextual menu pattern was again solved by triggering elements. The same concept was used to realise pop-up windows - sticky objects remaining to the front of the user, but with different background sound. The menu and toolbar was realised as shown in the example described above.

All sounds were audible when the user moved into their range. This means that there was silence in the starting position, but moving towards a wall meant that one could hear the 5 menu items at once at different levels depending on the distance.

Interaction with the prototype was done by joystick and keyboard. To avoid confusion while navigating through the virtual environment, no relative movements are supported. Bringing the joystick to the starting position means moving to the centre of the room facing the front wall. Moving up, along the z-axis, in the environment was implemented using the throttle handle of the joystick. The localisation of different sound sources was improved by using a headtracker.

3.2. Implementation

The prototype was implemented using Pure Data (PD) by Miller Puckette. For simulating the virtual environment the Binaural Ambisonics extension developed at the Institute of Electronic Music and Acoustics Graz was used [29, 30]. With this extension efficient binaural rendering of sound sources was possible including a room model with early reflections and late reverberation.

The content of the harddisk was faked in order to keep complexity low and not to introduce additional problems for the prototype interacting with the operating system.

3.3. Test Design

The test was performed by a group of 15 test participants divided into two groups. Group S were seven students of the Graz University of Technology and one person already holding a masters degree. All of them were between 20 and 27 years old and had good experience with computers and Windows. Group B consisted of four persons who are totally blind and three persons with visual disabilities. The use of visual screens was only feasible for them using additional magnification software. Six participants in group B hold the ECDL (European Computer Driving License) and use a computer in their work being very experienced with Windows software. One member had little experience with computers, but was attending the ECDL course. In average Group B was a little older.

After instruction, the participants got 15 minutes of training time with the application, participants were given a list of 7 tasks to perform. The tasks involved finding out how many files are in a specific folder, finding the size of files, copying, moving and creating files or folders.

Throughout the test, different types of data were collected. On the one hand, the hierarchic structure of files and folders after the test is stored in a text file. Apart from that, two further lists report the whole test sequence, one list containing the movement of the joystick within the virtual room (x,y,z-coordinates, rotation around the z-axis) in a resolution of 50 ms, the other list reporting any action performed by the participants with a time index, so that both lists can be combined. With these two lists, the whole test performance of the participants can be reproduced and visualised. The list of reported events can also be used to compute the quantity of different events. Apart from that, all the experimental sessions were monitored by the test administrator via headphones who additionally took notes.

After the test, the participants had to answer two questionnaires. One concerning the individual background of the participants, the other trying to catch the subjective impression of the participants after the test.

3.4. Analysis

The three-dimensional layout of the virtual room with the different meanings of the four surrounding walls and the two-staged movement (ground-plane movement towards the walls, vertical movement for selection) proved to be sufficient to host the elements of a real-world application. On the ground floor, at least 20 items (5 on each wall) can be placed, not to mention the potential with regard to vertical placement. The thematic grouping of elements on the different walls was easy to memorise for the test users. The usability of the mappings of particular interaction patterns is different. While the menu structure was easy to use for most test participants, the representation of the folder hierarchy is in need of improvement: Hardly any test user had a clear overview of the file and folder structure. According to the participants the static grid layout was confusing because they lost track of the absolute position while navigating through the tree structure. This was confirmed by the fact that hardly anyone could re-construct the file structure correctly after the test.

Remarkable is the percentage of user operations on items which are not related to the task, which were performed by mistake. Figure 6 shows that group B (the blind user group) required slightly more file handling operations to fulfil the tasks than group S (normally sighted user group). Hardly any file or folder was selected by mistake, the percentage of selected files that can be assigned to one of the tasks lies between 97% and 98% for all groups. The same was observed with menu selections. Group B needed slightly more menu operations, but no single user chose a menu item not related to the task he was performing.

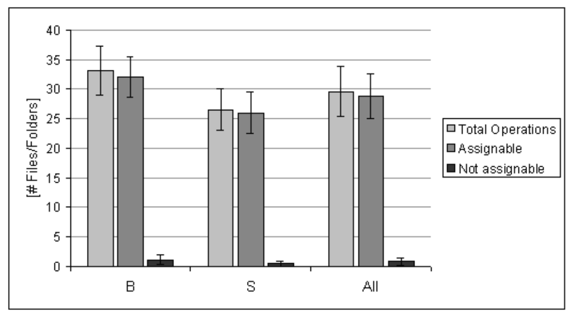

Figure 6: *Number of file/folder handling operations*

Throughout the whole test, the participants were free to choose how to fulfil the tasks: With the aid of the main menu, the standard buttons toolbar or the context menu. Figure 7 shows, which option was preferred by the two groups. Group B had a clear preference for the toolbar, whilst group S had a slight preference for the main menu. The usage of the context menu is similar for the group averages. The standard deviation is high for all values.

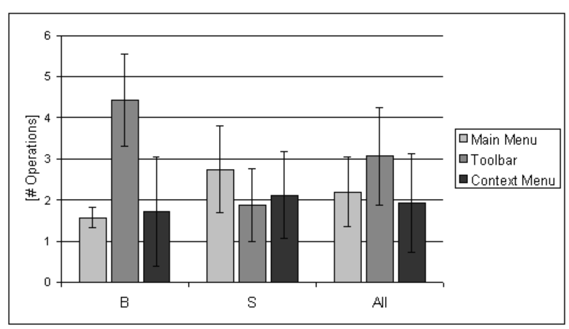

Figure 7: *Usage of different menus*

The virtual distance covered by the participants for one successful selection of an item shows that it lies within the dimension of the virtual room (10m x 14m and 12m high = 26m diagonal). Again it is slightly more for group B, but also lies within the standard deviation (Figure 8). These measures are comparable with

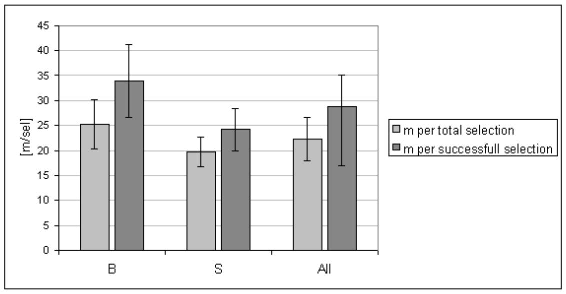

Figure 8: *Covered distance per selection*

mouse movements in graphical interfaces although no comparable data was collected for the graphical Explorer application.

In general, the subjective questionnaire showed that the users liked the system and felt comfortable solving the problems given. The sound design was chosen to be as non-obtrusive as possible and the silent centre position was appreciated.

The main goal of the application was to have equal usability for blind and sighted users. With regard to the main performance measurements (total test time, necessary file/folder handling operations, necessary menu handling operations, working speed and the percentage of correct results), the two groups of users obtained similar results, although there is a high standard deviation for some values.

4. CONCLUSIONS

This paper proposes a set of 20 mode independent interaction patterns for designing user interfaces in different interaction domains. The approach was chosen to overcome the difficulties of translating existing user interfaces into other domains like GUIs into auditory displays. The idea behind the patterns and their structure were explained followed by an evaluation of a real world application. The existing graphical user interface of this application was analysed, described through the proposed set of interaction patterns and transformed into the auditory domain.

The evaluation showed that the approach is feasible and promising for establishing design methodologies for auditory displays. Both groups of test participants, blind and normally sighted persons, were able to use the application with equal efficiency. All participants reported that they felt comfortable and could imagine operating a similar system in their real working environments.

Problems remained with some of the representations, where either sound design or the idea of representation was weak. However, having the user interface built from patterns and atoms, it is possible to isolate the problems and improve the transformation of the single item instead of re-designing the whole interface.

Future work needs to look closer into the mechanisms of acoustic communication considering psychoacoustics, learning effects, mental load and cross modal interactions. The patterns need to be revised and improved in their acoustical representation and more flexible frameworks must be developed in order to ease the integration of auditory displays into widely used computing systems. We also plan to look into formal descriptions of patterns, and develop tools to support both the design and implementation processes.

5. ACKNOWLEDGEMENTS

We would like to thank all who volunteered to participate in the evaluation test. Especially the people from the ISIS group (Integration, Service, Information and Schooling) in Graz / Austria for the close collaboration and for giving us an idea about the needs the visually impaired have when it comes to interaction with computers. ISIS is a project run by BFI (Berufsförderungsinstitut) Styria.

Support by the SonEnvir (`http://sonenvir.at`) project sponsored by the Zukunftsfonds Steiermark is gratefully acknowledged.

6. REFERENCES

[1] G. Kramer, Ed., *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading, MA: Addison-Wesley, 1994.

[2] C. Frauenberger, R. Höldrich, and A. de Campo, "A generic, semantically based design approach for spatial auditory computer displays," in *ICAD Proceedings*, Sydney, Australia, July 6–9 2004, Internation Conference on Auditory Display.

[3] T. Lokki, "Creating interactive virtual auditory environments," *IEEE Computer Graphics and Applications, special issue "Virtual Worlds, Real Sounds*, vol. 22, no. 4, pp. 49–57, July/August 2002, Electronic publication `http://www.computer.org/cga/`.

[4] J. S. Bamford, "An analysis of ambisonic sound systems of first and second order," M.S. thesis, University of Waterloo, `http://audiolab.uwaterloo.ca/~jeffb/thesis/thesis.html`, 1995.

[5] E. Verheijen, *Sound Reproduction by Wave Field Synthesis*, Ph.D. thesis, TU Delft, 1998.

[6] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, June 1997.

[7] M. Strauss, R. Höldrich, and A. Sontacchi, "A spatial audio interface for desktop applications," in *AES Proceedings, International Conference on Multichannel Audio*, Banff, Canada, June 26–28 2003, AES: Audio Engineering Society.

[8] C. Frauenberger, V. Putz, and R. Höldrich, "Spatial auditory displays - a study on the use of virtual audio environments as interfaces for users with visual disabilities," in *DAFx04 Proceedings*, Naples, Italy, October 5–8 2004, 7th Int. Conference on Digital Audio Effects (DAFx'04).

[9] D. K. McGookin and S. Brewster, "An investigation into the identification of concurrently presented earcons," in *ICAD Proceedings*, Boston, MA, USA, July 6–9 2003, International Community for Auditory Display, pp. 42–46.

[10] M. Gröhn, T. Lokki, and T. Takala, "Comparison of auditory, visual and audio-visual navigation in a 3d space," in *ICAD Proceedings*, Boston, MA, USA, July 6–9 2003, International Community for Auditory Display, pp. 200–2003.

[11] E. L. Oh and R. A. Lufti, "Informational masking by everyday sounds," *Journal of the Acoustical Society of America*, vol. 6, no. 106, pp. 3521–3528, 1999.

[12] M. Kobayashi and C. Schmandt, "Dynamic soundscape: Mapping time to space for audio browsing," in *ACM/SIGCHI 97 Proceedings*, Los Angeles, CA, March 22–27 1997, ACM Conference on Human Factors in Computing Systems, pp. 194–201.

[13] C. Schmandt, "Audio hallway: A virtual acoustic environment for browsing," in *Proc. ACM Conference of Computer Human Interactions*. April 18–23 1998, pp. 163–170, ACM Press, Los Angeles, California, USA.

[14] A. Walker, S. Brewster, and S.A. McGookin, "Diary in the sky: A spatial audio display for a mobile calendar," in *IHM-HCI Proceedings*, Lille, France, September 10–14 2001, Interaction Homme-Machine – Human Computer Interaction, IHM-HCI, pp. 531–540.

[15] W. K. Edwards, E. D. Mynatt, and T. Rodriguez, "The mercator project, a nonvisual interface to the x window system," *The X Resource*, 1993, O'Reilly Publishers.

[16] E. Mynatt, *Transforming Graphical Interfaces into Auditory Interfaces*, Ph.D. thesis, Georgia Institute of Technology, August 17 1995.

[17] M. Kaltenbrunner, "Y-windows: Proposal for a standard aui environment," in *ICAD Proceedings*, Kyoto, Japan, July 2–5 2002, International Community for Auditory Display.

[18] E. D. Mynatt and W. K. Edwards, *Extraordinary Human-Computer Interaction*, chapter Metaphors for Nonvisual Computing, Cambridge University Press, 1995.

[19] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human-Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989.

[20] S. Barrass, "Sonification design patterns," in *ICAD Proceedings*, Boston, USA, July 6–9 2003, Internation Conference on Auditory Display.

[21] M. Adcock and S. Barrass, "Cultivating design patterns for auditory displays," in *ICAD Proceedings*, Sydney, Australia, July 6–9 2004, Internation Conference on Auditory Display.

[22] K. V. Nesbitt, "Modeling the multi-sensory design space," in *CRPITS '01: Australian symposium on Information visualisation*, Darlinghurst, Australia, Australia, 2001, pp. 27–36, Australian Computer Society, Inc.

[23] J. Nielsen, *Usability Engineering*, Academic Press, London, 1993, ISBN 0125184050.

[24] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, and J. Neuhoff, "Sonification report: Status of the field and research agenda," `http://icad.org/websiteV2.0/References/nsf.html`, February 2005.

[25] A. D. Norman, *The Design of Everyday Things*, Doubleday Books, 1990.

[26] M. van Welie and H. Trætteberg, "Interaction patterns in user interfaces," in *PLOP Proceedings*, Monticello, Illinois, USA, August 13–16 2000, 7th. Pattern Languages of Programs Conference.

[27] V. Putz, "Spatial auditory user interfaces," M.S. thesis, Institute of Electronic Music and Acoustics, University of Music and Dramatic Arts Graz, 2004, `http://iem.at/projekte/dsp/spatial/dp_putz`.

[28] C. Frauenberger, V. Putz, R. Höldrich, and T. Stockma, "Mode independent interaction pattern design," in *Proceedings of 2nd International Conference on Non-visual & Multimodal Visualization*. IEEE Computer Society, 2005.

[29] M. Noisternig, T. Musil, A. Sontacchi, and R. Höldrich, "3d binaural sound reproduction using a virtual ambisonic approach," in *VECIMS Proceedings*, Lugano, Switzerland, 27–29 July 2003, International Symposium on Virtual Environments, Human-Computer Interfaces, and Measurement Systems, IEEE.

[30] M. Noisternig, R. Höldrich, T. Musil, and A. Sontacchi, "A 3d ambisonic based binaural sound reproduction system," in *AES 24, International Conference on Multichannel Audio*, Banff, Canada, June 26–28 2003, Audio Engineering Society.